UDC 534.86:
621.376.56

BBC

1969
THE QUEEN'S AWARD
TO INDUSTRY

RESEARCH DEPARTMENT

REPORT

PULSE-CODE MODULATION FOR HIGH-QUALITY SOUND-SIGNAL DISTRIBUTION: subjective effect of digit errors

No. 1972/40

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION

RESEARCH DEPARTMENT

PULSE-CODE MODULATION FOR HIGH-QUALITY SOUND-SIGNAL DISTRIBUTION: SUBJECTIVE EFFECT OF DIGIT ERRORS

Research Department Report No. **1972/40**
UDC 534.86:
621.376.56

This Report may not be reproduced in any form without the written permission of the British Broadcasting Corporation.

It uses SI units in accordance with B.S. document PD 5686.

G.R. Mitchell, T.Eng.(CEI), M.I.T.E.
M.E.B. Moffat, M.A., D.Phil., C.Eng., M.I.E.E., M.Inst.P.

P. Rainger
Head of Research Department

(EL-73)

Research Department Report No. 1972/40

PULSE-CODE MODULATION FOR HIGH-QUALITY SOUND-SIGNAL DISTRIBUTION: SUBJECTIVE EFFECT OF DIGIT ERRORS

Section Title Page

Summary 1

1. Introduction 1

2. Error-protection technique 1

3. Assessment of programme-quality impairment due to digit errors 2

3.1. Scope of subjective tests 2
3.2. Choice of test-programme material 2
3.3. Conduct of subjective tests 2
3.4. Results of the tests 3
    3.4.1. Programme-pause test (A1) 3
    3.4.2. Female-reader test (A2) 3
    3.4.3. Horse-race commentary intelligibility test (A3) 3
    3.4.4. Glockenspiel test (B1) 3
    3.4.5. String-quartet test (B2) 4
    3.4.6. Average results 5

4. Conclusions 5

5. References 6

6. Appendix 6

6.1. Details of test tapes 6
    6.1.1. Tests A1, A2, B1, B2: Tape recorded sequence 6
    6.1.2. Test A3: Tape recorded sequence 6
6.2. Details of test forms 6
    6.2.1. Series A form: test instructions 6
    6.2.2. Series B form: test instructions 7

(EL-73)

December 1972

Research Department Report No. 1972/40
UDC 534.86:
621.376.56

PULSE-CODE MODULATION FOR HIGH-QUALITY SOUND-SIGNAL DISTRIBUTION: SUBJECTIVE EFFECT OF DIGIT ERRORS

Summary

In a companion report (Report No. 1972/18) a digit-error-protection technique was recommended for use in a multiplex pulse-code modulation (p.c.m.) system for high-quality sound-signal distribution. The technique is briefly summarised in the present report. Tests have been carried out with and without the application of the protection technique to assess the impairment caused to critical programme material when random errors were introduced into the p.c.m. signal. It was found that the technique, as used in the recently commissioned BBC sound distribution links, afforded acceptable transmission performance at bit-error probabilities up to about 10^{-5}. *This represents a useful improvement over the performance with an unprotected channel.*

1. Introduction

For a recently completed study of multiplex distribution of high-quality sound signals by pulse-code modulation (p.c.m.),[1,2,3] a set of experimental equipment was constructed which incorporated a digit-error protection system as described in Section 2. The work described in this report was carried out to evaluate the effectiveness of the error-concealment technique employed.

Essentially, there were three parameters whose variations with bit-error probability were investigated. They were, first and second, subjective impairments of critical programme material with and without error concealment, respectively, and third, the intelligibility of speech. The investigation was confined to the subjective effects of random bit-errors, to which all transmission channels are prone. Other types of errors such as error bursts may sometimes be present as well, depending upon the characteristics of the particular bearer channel chosen for the p.c.m. signals. At the time that the work described here was carried out, it was known that the results of the study of multiplex p.c.m. mentioned above would probably be applied first to transmission over s.h.f. radio channels, but there was insufficient available knowledge of those channels to support burst-error tests. However, the results are valid for one kind of burst error, namely, that which is equivalent to periods of several seconds duration when a very high random-error rate prevails.

After describing the error-protection technqiue used, the report goes on to discuss the choice of programme material and the way in which tests were conducted to establish the subjective effect of bit errors upon that material.

2. Error-protection technique

A full description of earlier work on error-protection, and the application of the particular error-protection technique assessed in these tests, is discussed in another report.[3] The more important aspects of this technique may be summarised as follows.

In each sound channel the signal sample is converted into a binary number of 13 digits, or a 13-bit word, to which a check bit is added for error detection in the following manner. Among the group of six consecutive bits comprising the five most significant bits in the word and the check bit, an odd number of 'ones' is preserved, i.e. the check ('parity') bit is made one or zero to ensure that the number of ones among the six bits is odd. The parity bit is added before transmission, and is computed again at the receiving terminal. If the received and re-computed values of the parity bit are different, then a transmission error is indicated and an error-concealment technique is invoked in which the erroneous word is replaced by the preceding word. The latter technique is termed zero-order extrapolation. Although the technique described is an economic solution to the problem of random errors at low to moderate error probabilities (10^{-7} to 10^{-4}), it is desirable to mute the sound completely when the error probability is so high that the programme is unintelligible. This form of protection was not provided in the experimental equipment used here because speech intelligibility was one of the subjective factors being investigated.

The instrumental arrangement used for the tests described in this report was an experimental time-division multiplex p.c.m. system with one active channel, 12 channels bearing dummy data, and a digital synchronising signal ('framing pattern') whereby synchronisation between the sending and receiving terminals was maintained over the required range of bit-error probabilities (10^{-1} and below). The terminals were linked back-to-back via a co-axial cable, the errors being generated by the addition of white Gaussian noise at a tee-junction on the link. The average error rate in the active channel was measured by counting the detected errors at the output of the parity check unit, due account being taken of undetected errors in accordance with the curve of multiple bit-error probabilities given in Ref. 3, Fig. 4. For optimum error concealment[3] the sound-signal chain included high-frequency pre- and de-emphasis with a 50 μs characteristic, and a delay-line limiter to protect the analogue-to-digital converter from overload. The input programme level was reduced by 4 dB to avoid unacceptably frequent gain-reduction by the limiter. Pre- and de-emphasis was not included for the tests where the programme was unprotected from errors.

3. Assessment of programme-quality impairment due to digit errors

3.1. Scope of subjective tests

Only random bit-errors were introduced in the subjective tests. The bit-error probability was varied from 10^{-7} to 10^{-2} when evaluating subjective impairment with and without error protection in use, and from 3×10^{-3} to 10^{-1} for the tests in which speech intelligibility was assessed, again with and without error protection.

It was intended that the subjective tests at the lower error probabilities would show by how much the bit-error probability could be increased with error protection in use without introducing significant impairment to programmes. A reasonable threshold of acceptability of high-quality sound-signal distribution systems when expressed as an average subjective impairment on the EBU six-point scale, is grade 1·5. The impairment scale is shown in Table 1.

It was recognised that the speech intelligibility tests were not rigorous when judged by the standards of experts in the communication of marginally-intelligible speech. The intention was simply to establish a reasonably accurate value of bit-error probability at which listeners who were prepared to make an effort to understand a programme because they were keenly interested in its information content, as for example, with crucial stages of a horse-race commentary, reached a threshold of intelligibility. There would be little point in continuing with a programme if the bit-error probability exceeded this threshold value, and channel muting should then be applied.

TABLE 1

The EBU 6-Point Subjective Impairment Scale

Grade	Impairment
1	Imperceptible
2	Just perceptible
3	Definitely perceptible but not disturbing
4	Somewhat objectionable
5	Definitely objectionable
6	Unusable (Intolerable)

3.2. Choice of test-programme material

The choice of programme material for these tests was governed by three main factors:—

1. It was necessary for practical reasons to pre-record all of the test sequences on magnetic audio-tape. Programme signals which had been passed through the record-replay process only once or twice were essential to obtain the best signal-to-noise ratio and distortion level on the final presentation tapes. Masking of the effects of errors would thereby be minimised.

2. Several examples of critical material were desirable in order to cover the different subjective effects predominant at particular error rates. Errors in bits of lesser significance, which were always unconcealed, produced a programme-modulated noise at bit-error probabilities above about 10^{-5}, while occasional undetected word errors (two bit-errors in a word), significant only at probabilities above about 10^{-4}, caused loud clicks.

3. A previous preliminary investigation had indicated that a solo glockenspiel excerpt provided a very critical programme, and it was therefore used for this series of tests. A solo glockenspiel would however represent a rather rare or unusual programme and therefore the remaining excerpts were chosen, bearing in mind the other two factors, to be more representative of typical broadcast programmes.

The choice was further restricted by the need to minimise the duration of each subjective test to avoid listener fatigue, and five programme excerpts were finally chosen. Of these, one, the male-voice horse-race commentary, was used only at high error rates as an intelligibility test. The remaining four excerpts chosen were:— programme pause, female reader, glockenspiel, and string quartet. The three latter excerpts were each of approximately 30s duration. The programme pause consisted of 7s of a female voice relating the end of a short story followed by a 10s pause, followed in turn by 5s of glockenspiel, (representing a programme identification 'jingle'), the total duration being 22s.

The male-voice horse-race commentary excerpt used for the intelligibility test was 16s in duration. All of the items were halved in duration in the case of the higher error rates where a rapid assessment of grading was expected, in order to reduce listener fatigue.

3.3. Conduct of subjective tests

It was thought desirable that each test condition should be presented twice, to facilitate a rough check on the consistency of each listener's judgement. This, together with the need to cover adequately the relevant range of error rates, with and without error protection, resulted in a need for 24 presentations of each programme excerpt for each of the four impairment tests and 18 for the intelligibility test. Subjective test tapes were therefore prepared on this basis with the items presented in random order, as detailed in the Appendix.

The five tests were divided between two listening sessions as shown in Table 2.

TABLE 2

The Subjective Tests

First listening session (Subjective tests, series A)

Programme pause	A1
Female reader	A2
Male-voice horse-race commentary	A3

Total duration 29 minutes

Second listening session (Subjective tests, series B)

Glockenspiel	B1
String quartet	B2

Total duration 28 minutes

With the exception of the commentary test A3 for intelligibility, all tests were preceded by an example of unconcealed bit-errors alone, at a bit-error probability of 3×10^{-4}. In addition, the glockenspiel test B1 was preceded by a presentation of the excerpt unimpaired by bit-errors to assist the listeners in ignoring sounds which they might attribute to bit-errors or the error concealment technique.

In the four impairment tests, A1, A2, B1 and B2, the listeners were asked to grade the impairment of each of the 24 presentations using the EBU 6-point impairment scale (Table 1). For the programme-pause test, A1, the listeners were asked to assess the impairment during the pause. For the intelligibility test, A3, only two alternative grades were offered to the listeners for each of the 18 presentations, a 'U' if the commentary was considered unintelligible and an 'I' if intelligible. In all the tests each item was separated by a 5s period of silence, during which the listener was expected to write down his assessment. Details of the forms used in the tests are given in the Appendix, Section 6.2.

All of the tests were carried out in a listening room having acoustics similar to those of an average domestic living room. Not more than five listeners attended each session. The test recordings were reproduced at a sound-pressure level of approximately 85 dBA on a BBC monitoring loudspeaker type LS5/5. Nine observers participated in all of the tests, and two more took part in about half of them. All of the listeners were engineers experienced to some degree in this type of subjective test, and of these about half were concerned with the critical assessment of sound quality in their normal work.

3.4. Results of the tests

As stated in Section 3.3 each test condition was judged twice by the listeners. The average of the two assessments was used in the subsequent analysis of results.

Fig. 1 shows the mean subjective impairment grade plotted against bit-error probability for tests A1, A2, B1 and B2. Comparison of the graphs with and without error protection indicates the improvement due to the error-protection technique. The mean grade for nominally unimpaired items is quoted to indicate the degree to which the sound signals were impaired by the instrumentation in the absence of bit-errors.

3.4.1. Programme-pause test (A1)

Test A1 differed from all the others in that the listeners were asked to assess the impairment during a pause between two short programme excerpts. During the pause there were no programme-modulation effects due to the errors, no masking of errors by the programme, and, apart from normal equipment noise at a low level, instrumental imperfections were imperceptible. These facts probably explain why only in this test were the listeners unanimous in finding the impairment for the nominally unimpaired condition to be imperceptible.

As shown by Fig. 1(*a*), with error protection the bit-error probability corresponding to a grade of 1·5 was approximately 8×10^{-6} while that for grade 2 was approximately 2×10^{-5}. For a given bit-error probability up to 10^{-3}, error protection gave an improvement of at least 1·3 impairment grades.

3.4.2. Female-reader test (A2)

Test A2, inasmuch as the programme excerpt consisted of quiet speech with some short natural pauses, produced test conditions somewhat similar to those of the previous test. The results are given in Fig. 1(*b*). Comparing Figs. 1(*a*) and 1(*b*) it can be seen that the curves for error protection are similar, except that in Fig. 1(*b*) the linear portion is displaced slightly towards the region of higher error probability. This indicates some degree of error-masking by the programme. The listeners found the nominally unimpaired excerpt to be impaired with a mean grade of 1·2, due in part to the distortion introduced in the recording and replaying of the excerpt. With error protection the results corresponding to grades 1·5 and 2 were 10^{-5} and 3×10^{-5} respectively, and the average improvement was about 1·5 grades up to a bit-error probability of 10^{-3}.

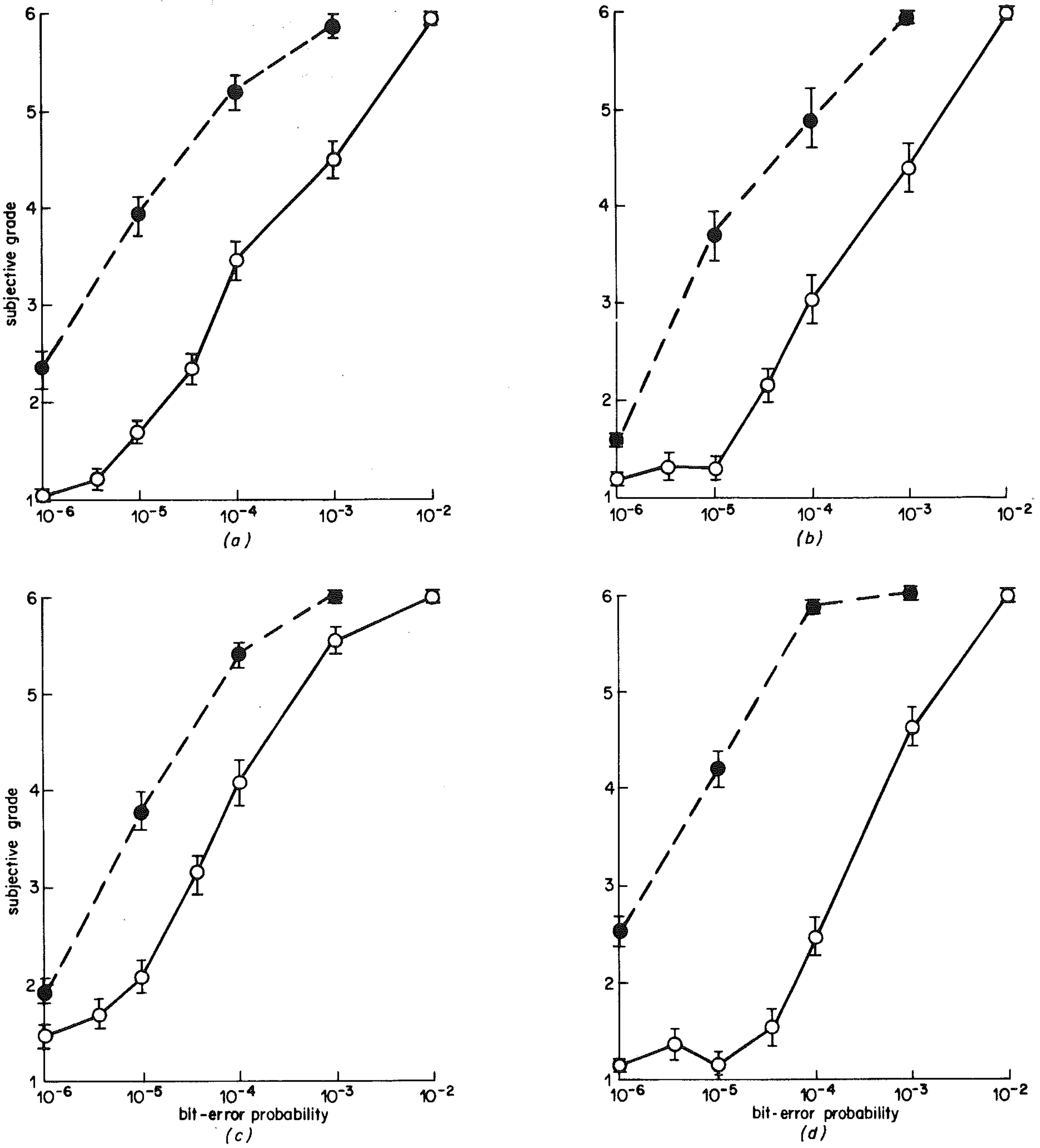
3.4.3. Horse-race commentary intelligibility test (A3)

The results of test A3 are plotted in Fig. 2 and show that the intelligibility threshold for 80% of the listeners was at a bit-error probability of about 5×10^{-2} with or without error protection.

It should be emphasised here, that for virtually all values of bit-error probability used in this test, the programme would have been graded in the normal way as unusable (grade 6). Only when the information in the programme is vital to a listener, e.g. the announcement of a national emergency, should he be expected to tolerate such a high error rate.

3.4.4. Glockenspiel test (B1)

The results for test B1 are plotted in Fig. 1(*c*). They confirm the glockenspiel excerpt as the most critical



(*a*) (*b*) (*c*) (*d*)

Fig. 1 - Mean subjective impairment grade versus bit-error probability

	Mean grade for nominally unimpaired items
(*a*) Programme-pause test, A1	1·0
(*b*) Female-reader test, A2	1·2
(*c*) Glockenspiel test, B1	1·5
(*d*) String-quartet test, B2	1·3

I Twice the standard error of the mean —○— Channel protected --●-- Channel unprotected

of those used, with error protection giving grade 1·5 at an error probability of 10^{-6} and grade 2 at approximately 10^{-5}. However, the relatively large impairment at 10^{-6} was almost certainly due to the fact that the listeners were influenced to a large degree by the nature of the programme and unavoidable imperfections in its recording and reproduction. The listeners gave a mean grade of 1·5 to the nominally unimpaired items in this test. Allowing for this it can be seen in Fig. 1(*c*) that error protection gave an improvement of at least one impairment grade up to a bit-error probability of about 2 x 10^{-4}.

3.4.5. String-quartet test (B2)

The results for test B2 are shown in Fig. 1(*d*) and indicate that, without error protection, we have a relatively

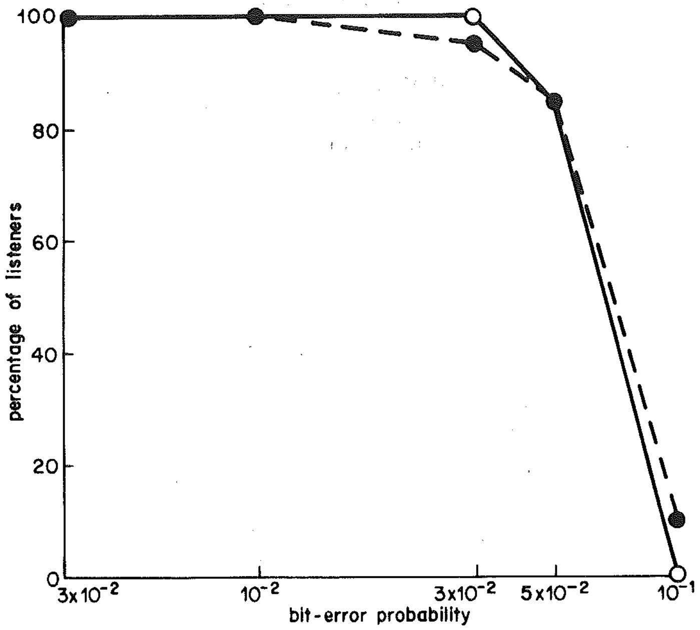


Fig. 2 - Horse-race commentary intelligibility test, A3 Percentage of listeners who found the excerpt intelligible, versus bit-error probability

—○— Channel protected – –●– – Channel unprotected

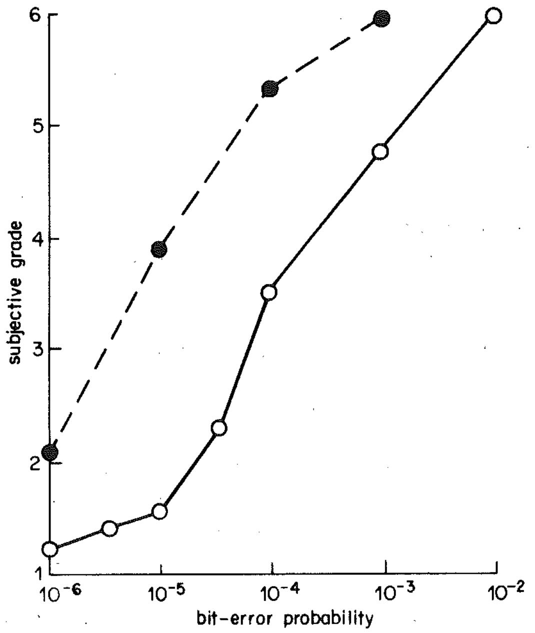


Fig. 3 - Average of the mean subjective impairments in tests A1, A2, B1, and B2, versus bit-error probability

Average of the mean grades for nominally unimpaired items = 1·25

—○— Channel protected – –●– – Channel unprotected

large impairment. The error protection technique used is, however, shown to be particularly effective for this kind of programme which is musically rich in harmonic and intermodulation content because of the non-linearities natural in string instruments. In this test grade 1·5 with error protection came at a bit-error probability of about 3 x 10^{-4}, and grade 2 at about 6 x 10^{-5}. Error protection gave an improvement of at least 1·5 impairment grades up to 10^{-3}; there was an improvement of 3 grades at around 10^{-5}. The mean grade for nominally unimpaired items was 1·3.

3.4.6. Average results

The averages of the mean results shown in Fig. 1 are shown in Fig. 3. The curves indicate that with error protection grade 1·5 is reached at a bit-error probability of about 10^{-5}, and grade 2 at about 2 x 10^{-5}. The improvement due to error protection is seen to be at least one impairment grade over most of the range tested; the maximum improvement was about 2·2 grades.

In view of the fact that the average of the mean grades for nominally unimpaired items is 1·25, it is likely that, but for distortion in the nominally unimpaired excerpts, the curve for error protection in Fig. 3 would pass through grade 1 rather than grade 1·2 at a bit-error probability of 10^{-6}.

4. Conclusions

The subjective effect of random bit-errors upon a selection of critical programme material has been investigated with and without the use of the error-protection technique recommended for use in the distribution of high-quality sound signals by p.c.m. Also, the threshold at which an attentive listener becomes unable to understand a speech programme has been evaluated in terms of random-bit-error probability.

Using the recommended error-protection technique, it was found that for the most critical programme item included in the tests, namely a solo glockenspiel, the bit-error probability could be as high as 10^{-6} without giving a subjective impairment worse than grade 1·5 on the EBU 6-point scale; grade 2 was reached at a bit-error probability of about 10^{-5}. The corresponding bit-error probabilities averaged for all the programme excerpts used in the tests were 10^{-5} and 2 x 10^{-5}, respectively. The use of the recommended error-protection technique was found to allow an increase by a factor of about 30 in the bit-error probability, for a given subjective grade, up to a bit-error probability of about 10^{-4}. Up to 10^{-3} the error protection technique gave an improvement of at least one impairment grade; the maximum improvement was about 2·5 grades at about 3 x 10^{-5}.

The threshold of male-speech intelligibility for 80% of the listeners was found to be at a bit-error probability of about 5 x 10^{-2} with or without error protection.

5. References

1. Pulse-code modulation for high-quality sound-signal distribution: feasibility of multiplex system. BBC Research Department Report No. 1970/34.

2. Pulse-code modulation for high-quality sound-signal distribution: instrumentation of experimental multiplex-system. BBC Research Department Report No. 1970/36.

3. Pulse-code modulation for high-quality sound-signal distribution: protection against digit errors. BBC Research Department Report No. 1972/18.

6. APPENDIX

6.1. Details of test tapes

6.1.1. Tests A1, A2, B1, B2: Tape-recorded sequence

Items in order on tape:–

400 Hz line up tone, 8 dB below peak–programme level.

(Test B1 only: example of programme excerpt unimpaired.)

Example of p.c.m. bit-errors alone without protection at bit-error probability of 3×10^{-4}.

Item	1	P at 10^{-4}
	2	U at 10^{-5}
	3	U at 10^{-4}
	4	P at 10^{-6}
	5	U at 10^{-3}
	6	P at 10^{-2}
	7	P at 3×10^{-5}
	8	U at 10^{-6}
	9	P at 3×10^{-6}
	10	P at 0
	11	P at 10^{-3}
	12	U at 10^{-4}
	13	P at 10^{-5}
	14	U at 10^{-3}
	15	P at 10^{-3}
	16	U at 10^{-6}
	17	P at 3×10^{-5}
	18	P at 0
	19	P at 3×10^{-6}
	20	P at 10^{-2}
	21	P at 10^{-6}
	22	U at 10^{-5}
	23	P at 10^{-4}
	24	P at 10^{-5}

U = Programme unprotected

P = Programme protected by single bit parity operating on the 5 most significant bits. With 50 μs pre- and de-emphasis and limiter the input programme level being reduced by 4 dB.

6.1.2. Test A3: Tape-recorded sequence

Items in order on tape:–

400 Hz line up tone, 8 dB below peak–programme level.

Item	1	U at 10^{-2}
	2	P at 5×10^{-2}
	3	U at 10^{-1}
	4	P at 10^{-2}
	5	U at 5×10^{-2}
	6	P at 3×10^{-2}
	7	U at 3×10^{-3}
	8	P at 10^{-1}
	9	U at 3×10^{-2}
	10	P at 5×10^{-2}
	11	P at 10^{-2}
	12	P at 3×10^{-2}
	13	U at 10^{-2}
	14	U at 10^{-1}
	15	U at 3×10^{-3}
	16	P at 10^{-1}
	17	U at 5×10^{-2}
	18	U at 3×10^{-2}

6.2. Details of test forms

For the subjective tests, forms were used on which boxes were provided for the entry of the listener's assessment of each item. The instructions given on the form for each test are reproduced below.

6.2.1. Series A form: test instructions

Test A1 (24 items: total duration 10 minutes):

An example of the audible effect of p.c.m. bit errors alone procedes the test. The test material simulates a short pause between two programmes, impaired by p.c.m. bit-errors. Grade the impairment of the pause caused by the p.c.m. bit-errors, bearing in mind the nature of the two adjacent programme excerpts. The EBU subjective impairment scale should be used. (Scale of Table 1 was given on the form.)

Test A2 (24 items: total duration 12·5 minutes):

The example of p.c.m. bit errors alone precedes the test. The test piece is an excerpt from the reading of a novel. Grade the impairment caused by the p.c.m. bit errors, again using the EBU scale.

Test A3 (18 items: total duration 6·25 minutes):

An excerpt from a horse-race commentary is seriously impaired by p.c.m. bit errors. State whether or not you would expect that an attentive and keenly-interested listener could understand the commentary sufficiently well to prevent him declaring the programme to be intolerable *because it is unintelligible.* In other words, our special listener should be reasonably confident that he has correctly understood the commentary.

Mark the box I for intelligible, U for unintelligible.

6.2.2. Series B form: test instructions

Test B1 (24 items: total duration 14·25 minutes):

An example of the test material, namely the playing of a glockenspiel, *unimpaired* by p.c.m. bit errors, and an example of the audible effect of p.c.m. bit errors alone, precede the test. Grade, according to the EBU scale shown below, the impairment of the test items caused by the p.c.m. bit errors. (Scale of Table 1 was given on the form.)

Test B2 (24 items: total duration 13·5 minutes):

The example of p.c.m. bit errors alone precedes the test. The test piece is an excerpt from the playing of a string quartet. Grade the impairment caused by the p.c.m. bit-errors, again using the EBU scale.

Printed by BBC Research Department, Kingswood Warren, Tadworth, Surrey